

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



554623

(43) Internationales Veröffentlichungsdatum
11. November 2004 (11.11.2004)

PCT

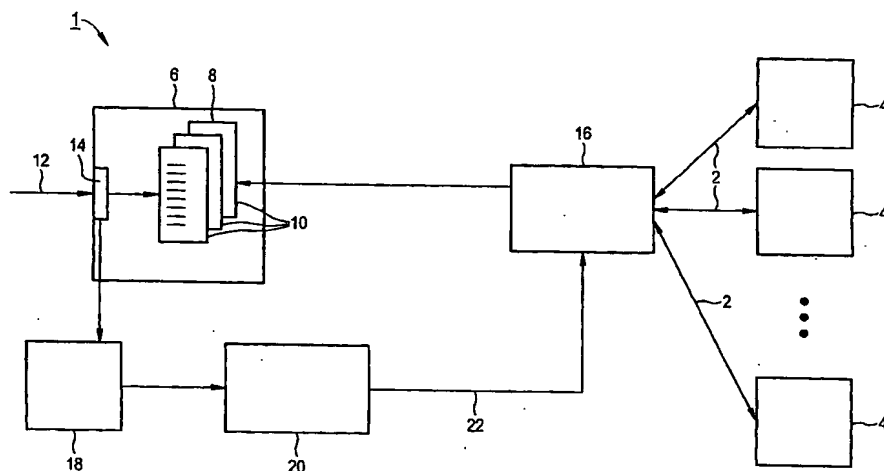
(10) Internationale Veröffentlichungsnummer
WO 2004/097670 A1

- (51) Internationale Patentklassifikation⁷: G06F 17/30
- (21) Internationales Aktenzeichen: PCT/EP2004/003972
- (22) Internationales Anmeldedatum:
15. April 2004 (15.04.2004)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
103 19 427.4 29. April 2003 (29.04.2003) DE
- (71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): CONTRACO CONSULTING & SOFTWARE LIMITED [GB/GB]; Athene House, The Broadway, London NW73 TB (GB).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): SCHAALE, Andreas [DE/DE]; Mulackstrasse 6, 10119 Berlin (DE).
- (74) Anwalt: WALKENHORST, Andreas; Tergau & Pohl, Eschersheimer Landstrasse 26, 60322 Frankfurt am Main (DE).
- (81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR GENERATING DATA RECORDS FROM A DATA BANK, ESPECIALLY FROM THE WORLD WIDE WEB, CHARACTERISTIC SHORT DATA RECORDS, METHOD FOR DETERMINING DATA RECORDS FROM A DATA BANK WHICH ARE RELEVANT FOR A PREDEFINED SEARCH QUERY AND SEARCH SYSTEM FOR IMPLEMENTING SAID METHOD

(54) Bezeichnung: VERFAHREN ZUR ERSTELLUNG VON FÜR DATENSÄTZE AUS EINER DATENBANK, INSBESONDERE AUS DEM WORLD-WIDE-WEB, CHARAKTERISTISCHEN KURZDATENSÄTZEN, VERFAHREN ZUR ERMITTLUNG VON FÜR EINE VORGEGBARE SUCHANFRAGE RELEVANTEN DATENSÄTZEN AUS EINER DATENBANK UND SUCHSYSTEM ZUR DURCHFÜHRUNG DES VERFAHRENS



(57) Abstract: The invention relates to a method for generating data records (4) from a data bank, especially from the World Wide Web, and characteristic short data records (10) for storing in a storage module (6) as a basis for determining the data records (4) that are relevant for a predefined search query. The aim of the invention is to provide a method that makes it possible to generate a search index that is particularly suitable for obtaining qualitatively high-value information from a data bank or from the World Wide Web. According to the invention, the system resources made available for generating a short data record (10) from a data record (4) are chosen taking into account the practical values obtained in previous search queries.

[Fortsetzung auf der nächsten Seite]

WO 2004/097670 A1



(84) **Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

— mit internationalem Recherchenbericht

— vor Ablauf der für Änderungen der Ansprüche geltenden Frist; Veröffentlichung wird wiederholt, falls Änderungen eintreffen

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(57) **Zusammenfassung:** Ein Verfahren zur Erstellung von für Datensätze (4) aus einer Datenbank, insbesondere aus dem World-Wide-Web, charakteristischen Kurzdatensätzen (10) zur Hinterlegung auf einem Speichermodul (6) als Basis zur Ermittlung der für eine vorgebbare Suchanfrage relevanten Datensätze (4) soll zur Erzeugung eines für eine qualitativ hochwertige Informationsbeschaffung aus der Datenbank oder aus dem World-Wide-Web besonders geeigneter Suchindex geeignet sein. Dazu werden erfindungsgemäss die zur Erstellung eines Kurzdatensatzes (10) aus einem Datensatz (4) bereitgestellten Systemressourcen unter Berücksichtigung von ermittelten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt.

Beschreibung

Verfahren zur Erstellung von für Datensätze aus einer Datenbank, insbesondere aus dem World-Wide-Web, charakteristischen Kurzdatensätzen, Verfahren zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen aus einer Datenbank und Suchsystem zur Durchführung des Verfahrens

Die Erfindung bezieht sich auf ein Verfahren zur Erstellung von für Datensätze aus einer Datenbank, insbesondere aus dem World-Wide-Web, charakteristischen Kurzdatensätzen zur Hinterlegung auf einem Speichermodul als Basis zur Ermittlung der für eine vorgebbare Suchanfrage relevanten Datensätze. Sie betrifft weiter ein Verfahren zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen aus einer Datenbank, insbesondere aus dem World-Wide-Web, bei dem derartige Kurzdatensätze auf ihre Relevanz für die jeweilige Suchanfrage hin durchsucht werden. Weiterhin betrifft die Erfindung ein Suchsystem zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen aus einer Datenbank, insbesondere aus dem World-Wide-Web.

In komplexen Datenbanken oder auch im weltumspannenden Computernetz („World-Wide-Web“) wird eine enorme Menge an Informationen bereitgehalten, die ein Benutzer zu Recherchezwecken mehr oder weniger gezielt abrufen kann. Um dabei Informationen aus der großen Informationsmenge heraus gezielt nutzbar machen zu können, kommen so genannte Suchmaschinen zum Einsatz, die gerade bei der Informationsbeschaffung aus dem World-Wide-Web teilweise erhebliche Verbreitung erfahren haben. Bei derartigen Suchmaschinen wird dem Benutzer üblicherweise über ein Ein-/Ausgabemodul ein Anfragenfenster zur Verfügung gestellt, über das gezielte Such- oder Recherchebegriffe vorgegeben werden können. Anschließend durchsucht die Suchmaschine die Informationsbasis der Datenbank bzw. des World-Wide-Web nach geeigneten Schlüssel- oder Keywörtern. Die daraufhin aufgefundenen Antwortdatensätze werden üblicherweise von der jeweiligen Suchmaschine hinsichtlich ihrer Relevanz für den vorgegebenen Suchauftrag kategorisiert und dem Benutzer in der Art einer Trefferliste in nach ihrer Relevanz geordneter Reihenfolge zur Verfügung gestellt.

Allerdings bereitet die zunehmende Komplexität der eingesetzten Datenbanken und insbesondere der stetig wachsende enorme Informationsgehalt im World-Wide-Web zunehmende Probleme bei der strukturierten und gezielten Beschaffung von Informationen. Die für die Recherche eingesetzten Suchmaschinen werden daher zunehmend
5 bezüglich der eingesetzten Suchalgorithmen verbessert, wobei zudem in der Art von Vorsortierungen oder Vorfilterungen weitere Hilfsmittel zur Klassifizierung von Datensätzen aus der Datenbank zum Einsatz kommen können. Insbesondere im World-Wide-Web sind die Datensätze üblicherweise in Form von so genannten Domains strukturiert und organisiert, wobei eine Domain typischerweise von einem Betreiber
10 unterhalten wird und ihrerseits eine Vielzahl von Unterdatensätzen, Textdokumenten oder dergleichen umfassen kann.

Um gerade bei der Informationsbeschaffung aus dem World-Wide-Web trotz der enormen Vielzahl der dort verfügbaren Datensätze oder Domains mit einem seinerseits
15 möglicherweise vergleichsweise großen Inhalt an Unterdatensätzen oder Informationsträgern eine geeignete Vorauswahl von bei einer vorgegebenen Suchanfrage zu berücksichtigenden Domains treffen zu können, kann ein so genanntes Ranking der Domains zum Einsatz kommen. Dabei wird jeder Domain ein Kennwert zugeordnet, der anhand von zugänglichen Sekundärinformationen in der Art einer relativen
20 Relevanz die Wichtigkeit der Berücksichtigung der jeweiligen Domain für die Suchanfrage charakterisiert. Bei der Zuordnung dieses Kennwerts zur jeweiligen Domain wird üblicherweise in der Art eines so genannten statischen Ansatzes eine Informationsbasis herangezogen, bei der beispielsweise aufgrund des Grades der Vernetzung der jeweiligen Domain mit anderen Domains auf die relative Bedeutung der jeweiligen Do-
25 main geschlossen wird. Als Maß für eine derartige Bedeutung kann die Anzahl der so genannten Links oder Querverweise von anderen Domains auf die jeweilige Domain herangezogen werden, wobei die Annahme zugrunde gelegt wird, dass eine Vielzahl von Querverweisen auf die jeweilige Domain ein Hinweis darauf ist, dass diese Domain für eine Vielzahl von Benutzern von besonderer Wichtigkeit bei der Bearbeitung ihrer
30 Suchanfragen ist.

Allerdings hat sich gezeigt, dass bei der Zuordnung einer derartigen statischen Kenngröße als Relevanzindikator für die jeweilige Domain Spielraum für Manipulationen

besteht, wobei unabhängig von den eigentlichen Benutzerinteressen aufgrund wirtschaftlicher Entscheidungen durch die Herstellung einer Vielzahl sachlich eigentlich nicht gerechtfertigter Links oder Querverweise künstlich der Eindruck einer vergleichsweise hohen Relevanz oder Bedeutsamkeit der jeweiligen Domain erzeugt werden kann. Die Verwendung derartiger statischer Relevanzzuordnungen zur Verbesserung der Suchergebnisse von Internet-Recherchen ist daher von eher abnehmendem Interesse.

Die enormen, im World-Wide-Web oder Internet verfügbaren Informationsmengen machen bei der Durchführung einer Informationsrecherche die tatsächliche Durchsuchung sämtlicher Domains inklusive der darin enthaltenen Unterdatensätze oder Textbausteine in Echtzeit auf das Vorhandensein der Suchanfrage oder von einzelnen Elementen der jeweiligen Suchanfrage unmöglich. Statt dessen kommen in Suchsystemen oder Suchmaschinen für die Informationsbeschaffung aus dem Internet oder World-Wide-Web so genannte „Crawler“ oder Durchsuchermodule zum Einsatz, die ständig die Domains oder Datensätze aus dem World-Wide-Web bzw. einer komplexen Datenbank auf ihren Textinhalt oder andere als relevant erachtete Informationen hin durchsuchen. Im Rahmen von vorgegebenen Systemressourcen (beispielsweise Bearbeitungszeit, Speicherkapazität oder Rechenleistung) durchsucht das jeweilige Durchsuchermodul die gerade angewählte Domain oder den Datensatz und stellt bis zu einer durch die zugewiesenen Systemressourcen vorgegebenen Grenze anhand der in der jeweiligen Domain aufgefundenen Informationen einen für die Domain oder den Datensatz charakteristischen Kurzdatensatz, beispielsweise in Form einer Textdatei mit möglicherweise zugeordneten Überschriften oder sonstigen Indikatoren, zusammen.

Dieser Kurzdatensatz wird sodann auf einem Speichermodul hinterlegt und für eine nachfolgende Untersuchung bereitgehalten. Die Gesamtheit der aus den überhaupt bei diesem Vorgehen berücksichtigten Datensätze oder Domains erstellten, auf dem Speichermodul hinterlegten Kurzdatensätze wird auch als so genannter „Index“ der jeweiligen Suchmaschine bezeichnet und dient als Informationsbasis für die nachfolgend durchgeführten Recherchen. Die Erstellung der im Index enthaltenen Kurzdatensätze erfolgt dabei üblicherweise kontinuierlich, wobei individuelle Domains zyklisch angewählt werden, so dass eine andauernde Aktualisierung des Index erfolgt. Bei einer

nachfolgenden Recherche, also bei der Ermittlung der für eine vorgegebene Suchanfrage relevanten Datensätze, wird sodann der durch die Gesamtheit der hinterlegten Kurzdatensätze gebildete Index auf das Vorhandensein von Schlüssel- oder Keywörtern der jeweiligen Suchanfrage oder von Einzelelementen davon durchsucht, wobei
5 anhand der dabei erzielten Suchergebnisse oder Treffer die den aufgefundenen Kurzdatensätzen zugeordneten Datensätze oder Domains als relevant für die jeweilige Suchanfrage ermittelt werden.

Aufgrund der Vielzahl der im World-Wide-Web verfügbaren Domains oder Datensätze
10 können nicht sämtliche Domains bei der Erstellung der Kurzdatensätze berücksichtigt werden. Die Entscheidung, welche Domains bei der Erstellung des Index berücksichtigt werden, erfolgt dabei üblicherweise anhand der oben genannten Relevanzkriterien, also insbesondere anhand von Informationen über eine erkannte oder vermeintliche Bedeutung der jeweiligen Domain für die Benutzer. Gerade da infolge der verfügbaren
15 Informationsvielfalt eine besonders zielgerichtete Vorsortierung der Informationen und insbesondere der als für eine Suchanfrage relevant erkannten Datensätze für die nachfolgende Auswertung der Rechercheergebnisse sehr bedeutsam sein kann, ist bereits bei der Erstellung des so genannten Index eine besonders hohe Qualität und Sorgfalt bei der Auswertung der berücksichtigten Informationen wünschenswert.

Der Erfindung liegt daher die Aufgabe zugrunde, ein Verfahren zur Erstellung von für
20 Datensätze charakteristischen Kurzdatensätzen der oben genannten Art anzugeben, mit dem ein für eine qualitativ hochwertige Informationsbeschaffung aus der Datenbank oder aus dem World-Wide-Web besonders geeigneter Suchindex erzeugt werden
25 kann. Des Weiteren sollen unter Verwendung dieses Verfahrens ein besonders geeignetes Verfahren zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen aus einer Datenbank, insbesondere aus dem World-Wide-Web, und ein Suchsystem zur Durchführung dieses Verfahrens angegeben werden.

30 Bezüglich des Verfahrens zur Erstellung der für die Datensätze charakteristischen Kurzdatensätze wird diese Aufgabe erfindungsgemäß gelöst, indem die zur Erstellung eines Kurzdatensatzes aus einem Datensatz bereitgestellten Systemressourcen unter

Berücksichtigung von ermittelten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt werden.

Die Erfindung geht dabei von der Überlegung aus, dass für die Erzeugung einer für
5 qualitativ besonders hochwertige Informationsbeschaffung besonders geeigneten Informationsbasis auf der Grundlage der für die Datensätze charakteristischen Kurzdatensätze zwar einerseits in der Art statischer Kennwerte verfügbare Informationen über die einzelnen Datensätze oder Domains an sich berücksichtigt werden können, das
10 aber andererseits in der Art eines dynamischen Elements auch für die Benutzerinteressen charakteristische Informationen mitberücksichtigt werden sollten. Dabei liegt die Erkenntnis zugrunde, dass das Ergebnis einer Informationsbeschaffung aus der Datenbank oder dem World-Wide-Web dann als qualitativ besonders hochwertig angesehen wird, wenn es in möglichst weit reichendem Umfang das Benutzerinteresse korrekt widerspiegelt. Es sollten daher Maßnahmen getroffen werden, um für das Benutzerinteresse
15 charakteristische Informationen in die weitere Informationsbeschaffung einfließen zu lassen. Ein Ansatzpunkt hierfür ist bereits die Erzeugung der Informationsbasis für die Bearbeitung der Suchanfragen, so dass sich Informationen über das Besucherinteresse bereits bei der Erzeugung der für die Datensätze oder Domains charakteristischen Kurzdatensätze im Index niederschlagen sollten. Um dies zu er-
20 möglichen, werden bereits bei der Zuweisung der Systemressourcen, die für die Erstellung eines Kurzdatensatzes aus einem zugeordneten Datensatz aufgewendet werden können und somit die Vollständigkeit der im Kurzdatensatz für die Informationsbeschaffung bereitgehaltenen Informationen entscheidend mitbestimmen, im Hinblick auf das Benutzerinteresse unter Berücksichtigung von Erfahrungswerten aus vorange-
25 gangenen Suchanfragen zugewiesen.

In besonders einfacher und prägnanter Weise können die Benutzerinteressen bereits bei der Erstellung des Index berücksichtigt werden, indem vorteilhafterweise bei der Zuordnung der Systemressourcen die Häufigkeit von zu einer Suchanfrage gleichen
30 oder ähnlichen Suchanfragen in der jüngsten Vergangenheit als Erfahrungswert berücksichtigt wird. In weiterer vorteilhafter Ausgestaltung kann dabei auch die Trefferhäufigkeit der Datensätze oder Domains bzgl. der von den Benutzern in jüngster Vergangenheit besonders häufig vorgegebenen Suchanfragen berücksichtigt werden. Die

Erfahrungswerte umfassen daher zweckmäßigerweise eine für die Anzahl gleichartiger Suchanfragen in einem vorgebbaren Zeitintervall charakteristische Kennzahl.

Um die Benutzerinteressen besonders zielführend bereits bei der Erstellung des Index für die Suchmaschine berücksichtigen zu können, werden vorteilhafterweise als Systemressourcen die Ressourcen eines zur Erstellung der für die Datensätze jeweils charakteristischen Kurzdatensätze vorgesehenen Durchsuchermoduls oder Crawlers unter Berücksichtigung von ermittelten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt.

In einer besonders vorteilhaften Ausgestaltung werden die Benutzerinteressen bei der Zuordnung der Systemressourcen besonders weitgehend berücksichtigt, indem bei der Ermittlung der Erfahrungswerte in besonderem Maß die möglicherweise komplexe Struktur der von den Benutzern verwendeten Suchanfragen berücksichtigt wird. Dabei liegt die Erkenntnis zugrunde, dass ein besonders genaues Abbild des allgemeinen Benutzerinteresses nicht allein durch die relative Häufigkeit einzelner in Suchanfragen verwendeter Elemente oder Begriffe, sondern ergänzend oder zusätzlich auch noch durch die Berücksichtigung von spezifischen Korrelationen zwischen einzelnen Begriffen oder Elementen von Suchanfragen erzielt werden kann. Dabei wird insbesondere berücksichtigt, dass einzelne Elemente oder Komponenten einer Suchanfrage entsprechend dem aktuell gerade verbreiteten Benutzerinteresse bevorzugt in Kombination mit spezifischen anderen Einzelelementen oder Komponenten von Suchanfragen angefragt werden. Beispielsweise könnte das aktuelle Benutzerinteresse im Allgemeinen aktuell in eine Richtung gehen, dass bevorzugt kostenlose Multimedia-Dateien aus dem Internet heruntergeladen werden sollen. In einer derartigen Konstellation ist bei Suchanfragen vermehrt mit einer Kombination der Suchbegriffe „MP3“, „free“ und „download“ zu rechnen. Bei der spezifischen Bewertung und Berücksichtigung vergangener Suchanfragen kann somit die Kombination dieser drei Einzelelemente einer Suchanfrage als ein besonders gewichtiger Indikator für gesteigertes Benutzerinteresse gewertet werden. Um dies zu ermöglichen, werden bei der Ermittlung der Erfahrungswerte vorzugsweise Korrelationen zwischen Einzelelementen der Suchanfragen berücksichtigt.

Um in der Art einer ersten Vorfilterung eine vergleichsweise einfach beschaffbare Eingangsinformation für die Bewertung von Suchanfragen und ihrer Relevanz für die Datensätze bereitzustellen, wird bei der Ermittlung der Erfahrungswerte vorteilhafterweise die relative Häufigkeit von Suchanfragen und/oder von Einzelelementen der Suchanfragen berücksichtigt. Dies kann auf besonders einfache Weise unmittelbar bei der Erstellung des Index berücksichtigt werden, indem denjenigen Datensätzen, die für eine vorgegebene Suchanfrage oder für eine vorgegebene Kombination aus Einzelelementen von Suchanfragen als relevant erkannt werden, vorteilhafterweise in Abhängigkeit von der relativen Häufigkeit der Suchanfrage bzw. der Kombination aus Einzelelementen von Suchanfragen zusätzliche Systemressourcen für die Erstellung des zugeordneten Kurzdatensatzes zugewiesen werden.

Vorteilhafterweise werden die auf die genannte Art erstellten, für die Datensätze aus der Datenbank charakteristischen Kurzdatensätze zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen aus der Datenbank, insbesondere aus dem World-Wide-Web, genutzt, indem die solchermaßen erstellten, in einem Speichermodul hinterlegten Kurzdatensätze auf ihre Relevanz für die jeweilige Suchanfrage hin durchsucht werden. Als Kriterium für die Ermittlung dieser Relevanz kann dabei beispielsweise die Häufigkeit dienen, mit der ein Schlüssel- oder Keyword der Suchanfrage im jeweiligen Kurzdatensatz aufgefunden werden kann, wobei zudem noch eine Differenzierung nach dem Ort des jeweiligen Auffindens, beispielsweise in einer Überschrift oder im Volltext, vorgenommen werden kann.

Bezüglich des Suchsystems zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen aus einer Datenbank, insbesondere aus dem World-Wide-Web, wird die genannte Aufgabe gelöst, indem in einem Speichermodul für die Datensätze charakteristische Kurzdatensätze hinterlegt sind, wobei die zur Erstellung eines Kurzdatensatzes aus einem Datensatz bereitgestellten Systemressourcen unter Berücksichtigung von hinterlegten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt sind.

Die Erfahrungswerte umfassen dabei vorteilhafterweise eine für die Anzahl gleichartiger Suchanfragen in einem vorgebbaren Zeitintervall charakteristische Kennzahl. In

zusätzlicher oder alternativer vorteilhafter Ausgestaltung sind als Systemressourcen die Ressourcen eines zur Erstellung der für die Datensätze jeweils charakteristischen Kurzdatensätze vorgesehenen Durchsuchermoduls unter Berücksichtigung von hinterlegten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt.

5 Die mit der Erfindung erzielten Vorteile bestehen insbesondere darin, dass durch die Berücksichtigung von Erfahrungswerten aus vorangegangenen Suchanfragen bei der Vergabe der Systemressourcen bei der Erstellung des Index oder der für die Datensätze charakteristischen Kurzdatensätze bereits in einem besonders frühen Stadium,
10 nämlich in der Vorbereitungsphase einer Datenbank- oder Internet-Recherche, eine weit gehende Berücksichtigung des aktuell vorhandenen Benutzerinteresses ermöglicht ist. Gerade durch die Berücksichtigung des Benutzerinteresses ergänzend zu oder anstelle von bisher verwendeten Datenbank spezifischen Charakteristika wie beispielsweise der Häufigkeit der jeweiligen Querverweise ist eine vom Benutzer als
15 besonders hochwertig angesehene Informationsbeschaffung ermöglicht. Ein besonders spezifisches Abbild des Benutzerinteresses und somit eine besonders hohe Genauigkeit bei der Vergabe der Ressourcen ist dabei erreichbar durch die Berücksichtigung von Korrelationen zwischen einzelnen Elementen von Suchanfragen, wobei gerade besonders häufig verwendete Kombinationen spezifischer Einzelelemente und der
20 Rückschluss auf die mit derartigen kombinierten Suchanfragen als Ergebnis aufgefundenen Datensätze oder Domains eine in besonderem Maße an das Benutzerinteresse angepasste Treffergenerierung erwarten lässt.

Ein Ausführungsbeispiel der Erfindung wird anhand einer Zeichnung näher erläutert.
25 Darin zeigt die Figur schematisch ein Suchsystem zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen oder Domains aus dem World-Wide-Web.

Das Suchsystem 1 gemäß der Figur ist über die mittels der Doppelpfeile 2 angedeuteten Datenleitungen des Internets oder World-Wide-Web mit einer Vielzahl von Domains 4 verbunden, wobei jede Domain 4 ihrerseits typischerweise eine Vielzahl von
30 Unterdatensätzen, Textbausteinen, multimedialen Informationselementen oder dergleichen umfasst.

Aufgrund der großen Vielzahl an im World-Wide-Web verfügbaren Informationen ist das Suchsystem 1 für die Bearbeitung einer Suchanfrage nicht für die Durchsuchung der Domains 4 oder der in ihnen enthaltenen Informationsinhalte auf das Vorhandensein bestimmter Schlüsselwörter oder Keywörter, sondern statt dessen für die Durchsuchung eines in einem Speichermodul 6 hinterlegten so genannten Index 8 ausgelegt. Der Index 8 umfasst eine Vielzahl von Kurzdatensätzen 10, von denen jeder jeweils für einen Datensatz oder eine Domain 4 des World-Wide-Web charakteristisch ist. Jeder Kurzdatensatz 10 enthält dabei einen als relevant erkannten Teil des Informationsgehalts der jeweils zugeordneten Domain 4, wobei im Kurzdatensatz 10 insbesondere die in der jeweiligen Domain 4 enthaltene Textinformation wiedergegeben ist. Zur Bearbeitung einer Suchanfrage wird diese, wie durch den Pfeil 12 angedeutet, einem Ein-/Ausgabemodul 14 des Suchsystems 1 zugeführt, von wo aus anhand von für die Suchanfrage charakteristischen Schlüssel- oder Keywörtern eine Durchsuchung der Kurzdatensätze 10 gestartet wird. Abhängig von der Anzahl der Ergebnisse oder Treffer, mit denen das Vorhandensein von Schlüssel- oder Keywörtern in den Kurzdatensätzen 10 festgestellt wird, wird die mit dem jeweiligen Kurzdatensatz 10 korrespondierende Domain 4 als für die Suchanfrage relevant erkannt und die entsprechende Domainadresse dem Benutzer auf einer Ergebnisliste mitgeteilt.

Zur Erstellung der für die Domains 4 charakteristischen Kurzdatensätze 10, die in ihrer Gesamtheit den Index 8 bilden, umfasst das Suchsystem 1 ein auch als „Crawler“ bezeichnetes Durchsuchermodul 16. Das Durchsuchermodul 16 nimmt in regelmäßigen, vorzugsweise zyklischen Zeitintervallen Kontakt mit den jeweiligen Domains 4 auf und durchsucht diese hinsichtlich ihres Informationsgehalts. Dabei kann insbesondere vorgesehen sein, die auf der jeweiligen Domain 4 hinterlegte Textinformation zu erfassen und geeignet zu komprimieren. Art und Umfang der Analyse des Inhalts jeder Domain 4 durch das Durchsuchermodul 16 werden dabei durch die Vorgabe spezifischer Systemressourcen des Durchsuchermoduls 16 für die jeweilige Domain 4 festgelegt. Als Systemressourcen können dabei abhängig von der jeweiligen Domain 4 beispielsweise die für die Durchsuchung vorgesehene Zeitspanne, die eingesetzte Rechnerleistung und/oder zugewiesene Speicherkapazitäten vorgegeben sein. Dabei kann insbesondere auch vorgegeben sein, ob die jeweilige Domain 4 vom Durchsuchermodul 16 überhaupt angesprochen oder von vornherein ignoriert werden soll. An-

hand der bei der Durchsuchung ermittelten Informationsbasis für die jeweilige Domain 4 wird anschließend vom Durchsuchermodul 16 in der Art einer Kurzfassung der zugeordnete Kurzdatensatz 10 erstellt und als Bestandteil des Index 8 im Speichermodul 6 abgelegt.

Die Zuordnung der Systemressourcen für die Durchsuchung der jeweiligen Domain 4 kann beispielsweise in Abhängigkeit domainspezifischer Relevanzkennwerte erfolgen. Dabei können auch so genannte statische Relevanzkennwerte vorgesehen sein, die anhand vorgegebener Kriterien wie beispielsweise dem Vernetzungsgrad einer Domain 4 mit anderen Domains 4 charakterisieren, wie hoch der Akzeptanzgrad der jeweiligen Domain 4 bei den Benutzern ist. Anhand dieser Relevanzkennwerte kann festgelegt sein, ob eine Domain 4 bei der Durchsuchung überhaupt berücksichtigt wird, und falls ja, wie sorgfältig die Durchsuchung der jeweiligen Domain 4 bei der Erstellung des zugeordneten Kurzdatensatzes 10 erfolgen soll.

Darüber hinaus ist das Suchsystem 1 aber auch dafür ausgelegt, bei der Erstellung der Kurzdatensätze 10 Erfahrungswerte und Erkenntnisse aus den vorangegangenen Suchanfragen mit zu berücksichtigen und somit das darin wiedergegebene aktuelle Benutzerinteresse in besonderem Maße in die Erstellung oder zyklische Erneuerung des Index 8 miteinfließen zu lassen. Dazu ist dem Speichermodul 6 ein weiteres Speichermodul 18 zugeordnet, in dem in der Art eines Logbuchs die eingehenden Suchanfragen zur weiteren Auswertung abgelegt werden. Die Inhalte des Speichermoduls 18 werden einem Analysemodul 20 zugänglich gemacht, das die eingegangenen Suchanfragen einer Auswertung unterzieht und anhand der dabei gewonnenen Erkenntnisse eine Neuverteilung der Systemressourcen auf die zu berücksichtigenden Domains 4 beim nächsten Durchsuchungszyklus vornimmt. Die entsprechende Zuordnung der Systemressourcen übermittelt das Analysemodul 20, wie über den Pfeil 22 dargestellt, anschließend an das Durchsuchermodul 16.

Bei der Zuordnung der Systemressourcen berücksichtigt das Analysemodul 20 somit Erfahrungswerte aus vorangegangenen Suchanfragen. Dies kann beispielsweise dadurch geschehen, dass die Häufigkeit einer Suchanfrage oder eines Schlüssel- oder Keywords als Einzelelement einer Suchanfrage ermittelt wird, wobei bei häufig verwen-

deten Suchanfragen oder Einzelelementen von Suchanfragen auf eine derzeit vergleichsweise hohe Popularität bei den Benutzern geschlossen wird. Dementsprechend wird angenommen, dass die bei vergleichsweise populären Suchanfragen aufgefundenen und als relevant erkannten Datensätze oder Domains 4 in vergleichsweise hohem Maße das aktuelle Benutzerinteresse widerspiegeln. In dieser Ausführungsform kann das Analysemodul 20 somit denjenigen Domains 4, die als Ergebnis für vergleichsweise häufig verwendete Suchanfragen aufgeführt werden, einen entsprechend erhöhten Anteil an Systemressourcen bei der nächstfolgenden Durchsuchung durch das Durchsuchermodul 16 zuordnen.

Darüber hinaus ist das Suchsystem 1 aber auch dafür ausgelegt, vergleichsweise komplexe Strukturen im Profil der Suchanfragen bei der Zuordnung der Systemressourcen durch das Analysemodul 20 mitzubetrachten. Dabei werden bei der Ermittlung der Erfahrungswerte insbesondere auch Korrelationen zwischen Einzelelementen von Suchanfragen mitberücksichtigt. Falls beispielsweise festgestellt wird, dass einzelne Elemente oder Suchwörter in Suchanfragen besonders häufig mit bestimmten anderen Einzelelementen oder Suchwörtern kombiniert werden, so wird auf eine hohe intrinsische Korrelation zwischen diesen beiden Suchelementen geschlossen, so dass einerseits diejenigen Domains 4, in denen vollständige oder approximative Kombinationen aufgefunden werden, als besonders relevant erkannt werden, wobei andererseits bei der Auswertung der relativen Häufigkeiten einzelner Suchelemente auch die relativen Häufigkeiten der damit besonders korrelierten weiteren Suchelemente berücksichtigt werden können.

Zur statistischen Auswertung der Suchanfragen wird dabei im Analysemodul 20 eine Korrelationsmatrix erstellt, deren Matrixelemente ein quantitatives Maß für die Korrelation zwischen jeweils zwei Einzelelementen von Suchanfragen angeben. Als quantitatives Maß kann dabei insbesondere die relative Häufigkeit, mit der die beiden jeweiligen Einzelelemente von Suchanfragen miteinander in Kombination gefragt werden, vorgesehen sein. Diese Korrelationsmatrix wird anschließend durch eine Hauptachsentransformation diagonalisiert, wobei auf der Hauptdiagonalen der diagonalisierten Matrix die Eigenwerte der ursprünglichen Korrelationsmatrix angegeben sind. Bei dieser

Hauptachsentransformation werden zudem die Eigenvektoren der Korrelationsmatrix ermittelt.

Die Eigenwerte und Eigenvektoren der Korrelationsmatrix können anschließend für eine weitere Auswertung der Suchanfragen herangezogen werden. Diejenigen Eigenvektoren der Korrelationsmatrix, die einen vergleichsweise großen Eigenwert aufweisen, entsprechen dabei einem Mix von Einzelelementen von Suchanfragen, der gemäß der Linearkoeffizienten der Einzelelemente der Suchanfragen vergleichsweise häufig in typischen Suchanfragen vorkommt und somit in besonderem Maße das aktuelle Benutzerinteresse widerspiegelt. Es werden somit in einem nächsten Schritt diejenigen Eigenvektoren der Korrelationsmatrix ausgewählt, denen ein vergleichsweise großer Eigenwert zugeordnet ist. Durch die so ermittelten Eigenvektoren erhält man im Ergebnis einen Mix von Suchanfragen, die in der jeweiligen Kombination mit besonders hoher Wahrscheinlichkeit in jüngster Vergangenheit vorgekommen sind.

Mit den so ausgewählten, vergleichsweise großen Eigenwerten der Korrelationsmatrix zugeordneten „Eigenanfragen“ greift das Analysemodul 20 in der Art einer Testanfrage auf den Index 8 zu und ermittelt so für jede „Eigenanfrage“ die zu dieser Eigenanfrage als relevant erkannten Datensätze oder Domains 4. Da die solchermaßen ermittelten Domains 4 in besonderem Maße dem aktuellen Benutzerinteresse entsprechen, werden für diese Domains 4 die Systemressourcen bei der erneuten Durchsuchung des World-Wide-Web im Vergleich zum vorigen Durchlauf proportional erhöht. Dies kann beispielsweise durch eine Zuweisung eines Wichtungsfaktors bei der Bereitstellung der Systemressourcen für die jeweilige Domain 4 nach der Beziehung

$$R_{vTA}(D_k) = (1 + \alpha * \lambda_k^\beta) \alpha, \beta > 0$$

erfolgen, wobei λ_k der Eigenwert der zugehörigen Eigenanfrage D_k eine auf diese Eigenanfrage als Treffer angezeigte Domain 4 und α eine geeignet gewählte Konstante >0 sein kann.

Bezugszeichenliste

1	Suchsystem
2	Doppelpfeile
4	Domain
6	Speichermodul
8	Index
10	Kurzdatensätze
12	Pfeil
14	Ein-/Ausgabemodul
16	Durchsuchermodul
18	Speichermodul
20	Analysemodul
22	Pfeil

- 14 -

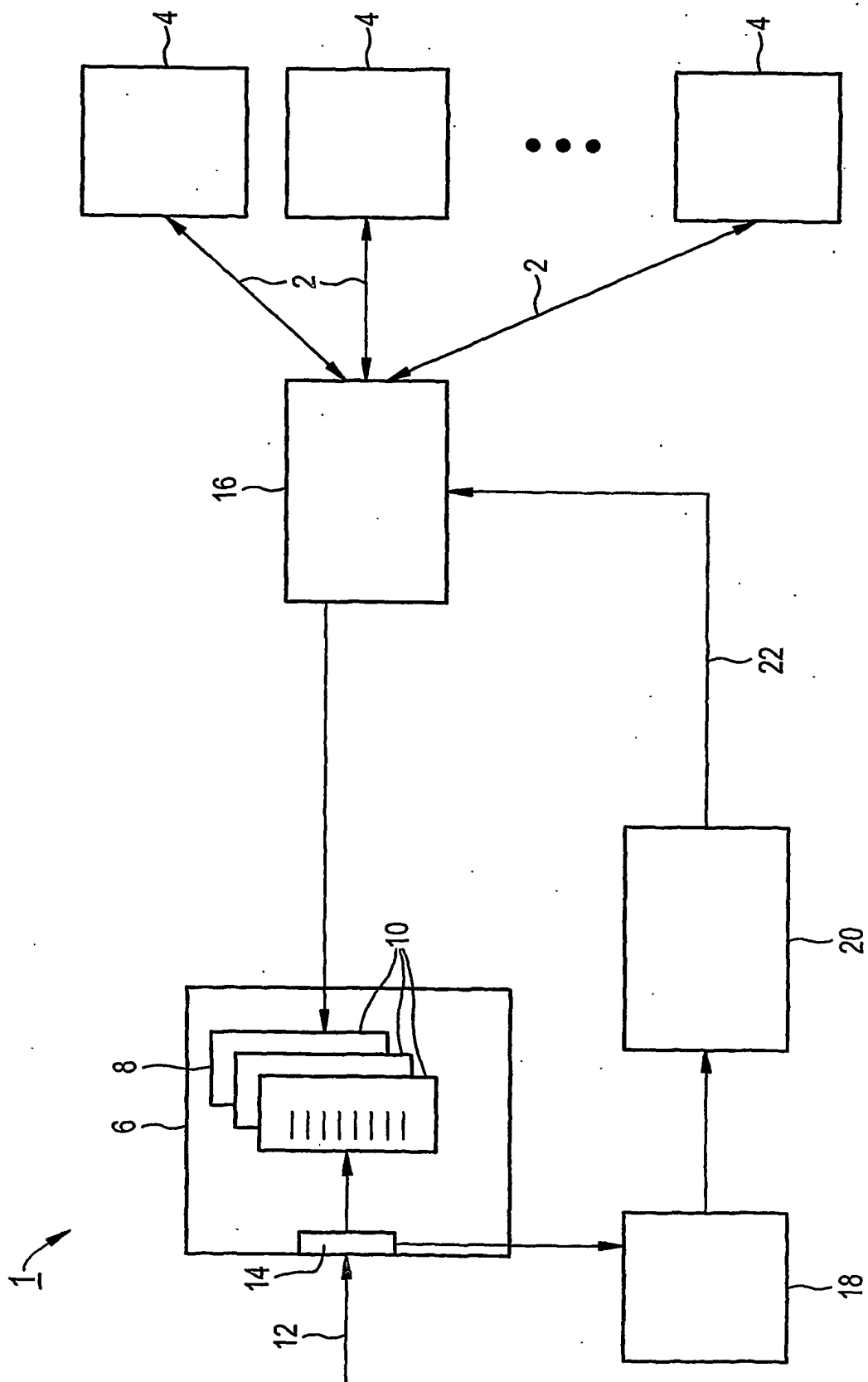
Ansprüche

1. Verfahren zur Erstellung von für Datensätze (4) aus einer Datenbank, insbesondere aus dem World-Wide-Web, charakteristischen Kurzdatensätzen (10) zur Hinterlegung auf einem Speichermodul (6) als Basis zur Ermittlung der für eine vorgebbare Suchanfrage relevanten Datensätze (4), bei dem die zur Erstellung eines Kurzdatensatzes (10) aus einem Datensatz (4) bereitgestellten Systemressourcen unter Berücksichtigung von ermittelten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt werden.
5
2. Verfahren nach Anspruch 1, bei dem die Erfahrungswerte eine für die Anzahl gleichartiger Suchanfragen in einem vorgebbaren Zeitintervall charakteristische Kennzahl umfassen.
15
3. Verfahren nach Anspruch 1 oder 2, bei dem als Systemressourcen die Ressourcen eines zur Erstellung der für die Datensätze (4) jeweils charakteristischen Kurzdatensätze (10) vorgesehenen Durchsuchermoduls (16) unter Berücksichtigung von ermittelten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt werden.
20
4. Verfahren nach einem der Ansprüche 1 bis 3, bei dem bei der Ermittlung der Erfahrungswerte Korrelationen zwischen Einzelelementen der Suchanfragen berücksichtigt werden.
25
5. Verfahren nach einem der Ansprüche 1 bis 4, bei dem bei der Ermittlung der Erfahrungswerte die relative Häufigkeit von Suchanfragen und/oder von Einzelelementen der Suchanfragen berücksichtigt wird.
- 30 6. Verfahren nach Anspruch 5, bei dem denjenigen Datensätzen (4), die für eine vorgegebene Suchanfrage oder für eine vorgegebene Kombination aus Einzelelementen von Suchanfragen als relevant erkannt werden, proportional zur relativen Häufigkeit der Suchanfrage bzw. der Kombination aus Einzelelementen von

Suchanfragen zusätzliche Systemressourcen für die Erstellung des zugeordneten Kurzdatensatzes (10) zugewiesen werden.

- 5 7. Verfahren zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen (4) aus einer Datenbank, insbesondere aus dem World-Wide-Web, bei dem in einem Speichermodul (6) hinterlegte, für die Datensätze (4) charakteristische, nach dem Verfahren nach einem der Ansprüche 1 bis 6 erstellte Kurzdatensätze (10) auf ihre Relevanz für die jeweilige Suchanfrage hin durchsucht werden.
- 10 8. Suchsystem (1) zur Ermittlung von für eine vorgebbare Suchanfrage relevanten Datensätzen (4) aus einer Datenbank, insbesondere aus dem World-Wide-Web, insbesondere zur Durchführung des Verfahrens nach Anspruch 7, bei dem in einem Speichermodul (6) für die Datensätze charakteristische Kurzdatensätze (10) hinterlegt sind, wobei die zur Erstellung eines Kurzdatensatzes (10) aus einem
15 Datensatz (4) bereitgestellten Systemressourcen unter Berücksichtigung von hinterlegten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt sind.
- 20 9. Suchsystem (1) nach Anspruch 8, bei dem die Erfahrungswerte eine für die Anzahl gleichartiger Suchanfragen in einem vorgebbaren Zeitintervall charakteristische Kennzahl umfassen.
- 25 10. Suchsystem (1) nach Anspruch 8 oder 9, bei dem als Systemressourcen die Ressourcen eines zur Erstellung der für die Datensätze (4) jeweils charakteristischen Kurzdatensätze (10) vorgesehenen Durchsuchermoduls (16) unter Berücksichtigung von hinterlegten Erfahrungswerten aus vorangegangenen Suchanfragen gewählt sind.

1/1



INTERNATIONAL SEARCH REPORT

International Application No

CT/EP2004/003972

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 6 493 703 B1 (GROSS J NICHOLAS ET AL) 10 December 2002 (2002-12-10) column 4, line 66 - column 5, line 16 column 6, line 40 - column 7, line 17 column 9, line 25 - column 10, line 39 column 19, line 24 - column 20, line 49	1-3,6-10 4,5
X	WO 02/08962 A (ENERGY E COMM COM INC) 31 January 2002 (2002-01-31) page 4, line 15 - page 4, line 28 page 9, line 17 - page 9, line 28 claims 1-7 ----- -/-	1-3,6-10

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- * & * document member of the same patent family

Date of the actual completion of the international search

16 September 2004

Date of mailing of the international search report

23/09/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Abbing, R

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ROCHA L M: "Adaptive Webs for Heterarchies with Diverse Communities of Users" WORKSHOP FROM INTELLIGENT NETWORKS TO THE GLOBAL BRAIN: EVOLUTIONARY SOCIAL ORGANIZATION THROUGH KNOWLEDGE TECHNOLOGY, XX, XX, 3 July 2001 (2001-07-03), pages 1-35, XP002209508	4,5
A	page 10, line 30 - page 17, line 3 page 19, line 9 - page 23, line 9 page 24, line 11 - page 25, line 2	1-3,6-10
A	US 2002/194161 A1 (HALL MARTIN R ET AL) 19 December 2002 (2002-12-19) page 1, paragraph 10 - page 1, paragraph 11 page 2, paragraph 20 - page 3, paragraph 25 page 3, paragraph 29 - page 3, paragraph 30	1-3,6-10
A	EP 1 207 468 A (ITT MFG ENTERPRISES INC) 22 May 2002 (2002-05-22) column 1, paragraph 8 - column 2, paragraph 11	1-3,6-10
A	US 6 418 433 B1 (CHAKRABARTI SOUMEN ET AL) 9 July 2002 (2002-07-09) abstract column 2, line 56 - column 3, line 30	1,8
A	ARASU A ET AL: "SEARCHING THE WEB" ACM TRANSACTIONS ON INTERNET TECHNOLOGY, ACM, NEW YORK, NY, US, vol. 1, no. 1, August 2001 (2001-08), pages 2-43, XP001143684 ISSN: 1049-3301 figure 1 page 3, line 38 - page 5, line 40 page 6, line 7 - page 8, line 13	1,8
A	WO 02/27562 A (NINESIGMA INC) 4 April 2002 (2002-04-04) page 3, line 13 - page 3, line 25 page 8, line 23 - page 10, line 30	1,8

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

CT/EP2004/003972

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
US 6493703	B1	10-12-2002	US 6721748 B1	13-04-2004
WO 0208962	A	31-01-2002	AU 7800401 A	05-02-2002
			WO 0208962 A1	31-01-2002
			US 2002069203 A1	06-06-2002
US 2002194161	A1	19-12-2002	NONE	
EP 1207468	A	22-05-2002	GB 2368935 A	15-05-2002
			EP 1207468 A2	22-05-2002
			US 2002059399 A1	16-05-2002
US 6418433	B1	09-07-2002	NONE	
WO 0227562	A	04-04-2002	US 6584468 B1	24-06-2003
			AU 9319301 A	08-04-2002
			WO 0227562 A2	04-04-2002

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
IPK 7 G06F17/30

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
IPK 7 G06F

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)
EPO-Internal, INSPEC

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X Y	US 6 493 703 B1 (GROSS J NICHOLAS ET AL) 10. Dezember 2002 (2002-12-10) Spalte 4, Zeile 66 - Spalte 5, Zeile 16 Spalte 6, Zeile 40 - Spalte 7, Zeile 17 Spalte 9, Zeile 25 - Spalte 10, Zeile 39 Spalte 19, Zeile 24 - Spalte 20, Zeile 49	1-3,6-10 4,5
X	WO 02/08962 A (ENERGY E COMM COM INC) 31. Januar 2002 (2002-01-31) Seite 4, Zeile 15 - Seite 4, Zeile 28 Seite 9, Zeile 17 - Seite 9, Zeile 28 Ansprüche 1-7 ----- -/-	1-3,6-10

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

☒ Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

A Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

E älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

L Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

O Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

P Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

T Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

X Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

Y Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

Z Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

16. September 2004

Absendedatum des internationalen Recherchenberichts

23/09/2004

Name und Postanschrift der Internationalen Recherchenbehörde
Europäisches Patentamt, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Abbing, R

C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
Y	ROCHA L M: "Adaptive Webs for Heterarchies with Diverse Communities of Users"	4,5
A	WORKSHOP FROM INTELLIGENT NETWORKS TO THE GLOBAL BRAIN: EVOLUTIONARY SOCIAL ORGANIZATION THROUGH KNOWLEDGE TECHNOLOGY, XX, XX, 3. Juli 2001 (2001-07-03), Seiten 1-35, XP002209508	
A	Seite 10, Zeile 30 - Seite 17, Zeile 3 Seite 19, Zeile 9 - Seite 23, Zeile 9 Seite 24, Zeile 11 - Seite 25, Zeile 2	1-3,6-10
A	US 2002/194161 A1 (HALL MARTIN R ET AL) 19. Dezember 2002 (2002-12-19) Seite 1, Absatz 10 - Seite 1, Absatz 11 Seite 2, Absatz 20 - Seite 3, Absatz 25 Seite 3, Absatz 29 - Seite 3, Absatz 30	1-3,6-10
A	EP 1 207 468 A (ITT MFG ENTERPRISES INC) 22. Mai 2002 (2002-05-22) Spalte 1, Absatz 8 - Spalte 2, Absatz 11	1-3,6-10
A	US 6 418 433 B1 (CHAKRABARTI SOUMEN ET AL) 9. Juli 2002 (2002-07-09) Zusammenfassung Spalte 2, Zeile 56 - Spalte 3, Zeile 30	1,8
A	ARASU A ET AL: "SEARCHING THE WEB" ACM TRANSACTIONS ON INTERNET TECHNOLOGY, ACM, NEW YORK, NY, US, Bd. 1, Nr. 1, August 2001 (2001-08), Seiten 2-43, XP001143684 ISSN: 1049-3301 Abbildung 1 Seite 3, Zeile 38 - Seite 5, Zeile 40 Seite 6, Zeile 7 - Seite 8, Zeile 13	1,8
A	WO 02/27562 A (NINESIGMA INC) 4. April 2002 (2002-04-04) Seite 3, Zeile 13 - Seite 3, Zeile 25 Seite 8, Zeile 23 - Seite 10, Zeile 30	1,8

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2004/003972

Im Recherchenbericht angeführtes Patentdokument		Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
US 6493703	B1	10-12-2002	US 6721748 B1	13-04-2004
WO 0208962	A	31-01-2002	AU 7800401 A	05-02-2002
			WO 0208962 A1	31-01-2002
			US 2002069203 A1	06-06-2002
US 2002194161	A1	19-12-2002	KEINE	
EP 1207468	A	22-05-2002	GB 2368935 A	15-05-2002
			EP 1207468 A2	22-05-2002
			US 2002059399 A1	16-05-2002
US 6418433	B1	09-07-2002	KEINE	
WO 0227562	A	04-04-2002	US 6584468 B1	24-06-2003
			AU 9319301 A	08-04-2002
			WO 0227562 A2	04-04-2002